

Introduction to Deep Learning: Part I

HONGBO ZHOU, Statoil, LASSE AMUNDSEN and MARTIN LANDRØ

Once, artificial intelligence (AI) was science fiction. Today, it is part of our everyday lives. Tomorrow, AI is speculated to make computers smarter than people, and perhaps threaten the survival of humankind. In the future, can computers begin to think for themselves? What are the trends in AI? What is to come?

Encyclopaedia Britannica defines Artificial Intelligence, or AI as it is commonly called, as the ability of a computer or computer-controlled robot to perform tasks that normally require human intelligence, such as the ability to reason, discover meaning, generalise, or learn from past experiences.

We have seen AI robots in movies or read about them in science fiction novels. C-3PO is a robot character from the Star Wars universe whose main function is to assist etiquette, customs, and translation, so that meetings of different cultures run smoothly. On the evil side, recall the ‘Terminator’ series. Before becoming self-aware, Skynet is a powerful AI system for the US military to coordinate national defence; after becoming self-aware, Skynet decides to coordinate the destruction of the human species instead, with the Terminator robots serving as its agents, disguised as humans.

Whilst the idea of AI can be terrifying, there are interesting ‘passive’ forms of real AI. First, however, we will look briefly into the history of AI.

Early AI Milestones

The earliest work in the field of AI was done in the mid-20th century by the British mathematician and computer pioneer Alan Turing. In 1947, he discussed computer

intelligence in a lecture, saying, “What we want is a machine that can learn from experience,” and that the “possibility of letting the machine alter its own instructions provides the

mechanism for this.” In 1950, he wrote a paper, ‘Computing machinery and intelligence,’ addressing the issue of AI.

One of the earliest successful demonstrations of the ability of AI

Artificial intelligence is the use of computers to simulate human intelligence. Deep learning – driven by ever more powerful GPUs – grows more useful as the amount of data in the world grows. This image is NVIDIA’s brand representation of their AI podcast, where experts discuss how AI works, how it is evolving, and how it is being used across industries.

“We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run.”

*Roy Amara (1925-2007)
American scientist, futurist and president
of the Institute for the Future.*



© 2017 NVIDIA Corporation. All rights reserved. Image provided courtesy of NVIDIA Corporation.

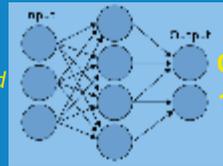
AI Progression

ARTIFICIAL INTELLIGENCE

Machine Learning

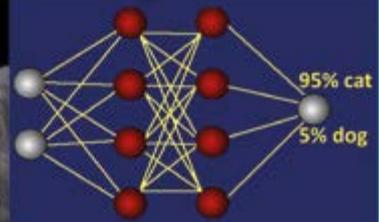


Date, time
Salary, m. spend
Loc. payment



0 – No Fraud
1 – Fraud

Deep Learning



1950s

1960s

1970s

1980s

1990s

2000s

2010s

Machine learning and deep learning are subsets of AI. One of the first AI programs was a checkers (or draughts) program, written by Arthur Samuel, which won a game against a former checkers champion in 1962. Machine learning is applied to automatically discover patterns in data, which can be used to make predictions. For instance, a machine learning system can learn patterns in credit card transactions in a bank's database that are predictive of fraud. The more data (e.g. date, time, salary, average

monthly spend, location, merchant, price and whether the transaction was legitimate or not) the system processes, the better its predictions become.

Deep learning is driving today's AI. It is widely used for tasks like face-tagging of photos, voice recognition, and language translation. There is now hope that deep learning will be able to diagnose deadly diseases and do countless other things to transform whole industries, including the oil and gas industry.

programs to incorporate learning was published in 1952. Anthony Oettinger at the University of Cambridge, influenced by the views of Alan Turing, developed the response learning program 'Shopper', in which the universe was a mall of eight shops. When sent out to purchase an item, Shopper would visit these shops at random until the item was found, but while searching it would memorise a few of the items stocked in each shop visited. The next time Shopper was instructed to get the same item, or some other item that it had already located, it would go to the right shop straight away. This simple form of learning is called rote learning, a memorisation technique

based on repetition without proper understanding or reflection. Today, we note that AI in online shopping is big business. AI technology allows businesses to analyse the customer's behaviour, predict consumer needs and offer tailored customer experiences. AI is designed to make online experiences altogether more personal.

The 1956 Dartmouth Artificial Intelligence Conference marked the birth of the field of AI as a vibrant area of interdisciplinary research; many of the attendees later became leaders in AI research. These pioneers were optimistic about the future and believed that within two decades machines would be capable of doing any work a person can do. Their

attitude was shown in their proposal: "a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer". Their dream was to construct complex machines – enabled by emerging computers – that possessed the characteristics of human intelligence.

After expressing the bold goal of simulating human intelligence, researchers developed a range of demonstration algorithms that showed that computers could perform tasks once thought to be solely the domain of human capability. However, lack of computer power soon stopped progress, and by the mid-1970s AI was considered overhyped and tossed into

Recent Advances in Technology

technology's trash heap. Technological work on AI had to continue with a lower profile.

Machine Learning – An Approach to Achieve AI

In the 1990s, machine learning, a subset of AI, started to gain popularity. The machine learning field changed its goal from achieving AI to tackling solvable problems of a practical nature. Machine learning adapted methods and models borrowed from statistics and probability theory. Among the most common methods are artificial neural networks or ANNs (weighted decision paths), which are electronic networks of 'neurons' loosely analogous to the neural structure of the brain, and genetic algorithms, which aim to evolve solutions to problems by iteratively generating candidate solutions, culling the weakest, and introducing new solution variants by introducing random mutations.

Machine learning thus has links to optimisation. Many learning problems can be formulated as minimisation of some objective or loss function on a training set of examples. Loss functions express the misfit between the predictions of the model being trained and the actual problem instances; for example, in classification, one wants to assign a label to instances, so models are trained to correctly predict the pre-assigned labels of a set of examples. The difference between optimisation and machine learning lies in their goals: while the goal of optimisation algorithms is to minimise the loss on a given training set, the goal of machine learning is the prediction of unseen samples. In this way, the machine learning discipline is concerned with the implementation of computer software that can learn autonomously.

Machine learning is mainly about feature extraction, i.e., the extraction of representations or abstractions that are pieces of information or characteristics that might be useful for prediction. Historically, there exist two major arenas of machine learning: the traditional computationalism concept that mental activity is computational and symbolic or logic-based; and

the connectionism view, in which mental activity can be described by interconnected networks of simple units and is neural-based. Neural networks, as we will discuss below, are by far the most commonly used connectionist model today. These two scenarios, however, have duelled each other since their birth.

Computation or Neural Networks

Traditional symbolic-based machine learning models depend heavily on feature engineering, a process of using domain knowledge to manually extract features that make machine learning algorithms work. Specifically, the programmer needs to tell the computer the kinds of things it should be looking for that will be informative in decision-making. The algorithm's effectiveness relies on how insightful the programmer is. For complex problems like object recognition, this proves to be both difficult and time-consuming, meaning that feeding the algorithm with raw data rarely ever works for traditional symbolic-based machine learning. But, unlike its rival, the ANNs system, people have full control of obtaining what they want to achieve.

Consider this example: a human driver uses his eyes and brain to see and visually sense the traffic around him. When he sees a red rectangular-shaped plate with a white border and large white letters saying WRONG WAY, he knows that if he drives pass the sign, he is in trouble. For many years experts tried to use machine learning to teach computers to recognise signs in the same way. The solution, however, required hand-coding. Programmers would write classifiers such as edge detection filters so the program could identify where an object started and stopped; shape detection routines to determine if the object had four sides; and a routine to recognise the letters 'Wrong Way'. From all those hand-coded classifiers they would develop a theoretical and algorithmic basis to achieve automatic visual understanding. But would you trust the computer if a tree obscures part of the sign?

The other arena for machine



learning is ANNs. Neural networks, based on learning multiple levels of representation or abstraction, have traditionally been viewed as simplified models of neural processing in the brain, even though the relation between this model and the biological architecture of the brain is debated as it is not clear to what degree ANNs mirror brain function. Over the past few decades computer scientists have developed various algorithms that try to allow computers to learn to solve problems automatically through Big Data. ANNs have been successful in various applications in recent years but the criticism remains about its opaqueness. People have some clues on how to make it work, but do not actually know why it works so well.

In the next issue of GEO ExPro we will extend our discussion on deep learning. ■

The MNIST dataset is a standard benchmark dataset for machine learning. It is a modified subset of two datasets collected by National Institute of Standards and Technology (NIST). It contains 70,000 scanned images of handwritten digits from 250 people, half of whom were US Census Bureau employees, the rest being high school students. There have been numerous attempts to achieve the lowest error rate in solving the handwritten digit recognition problem; one attempt, using a hierarchical system of convolutional neural networks, manages to get an error rate on the MNIST database of 0.23 %.

